Math 408 Homework Set 4 Solutions

(1) Let H, A, and b be as above and define $\ell : \mathbb{R}^k \to \mathbb{R}$ by

$$\ell(x) := \frac{1}{2}(Ax-b)^TQ(Ax-b) \ ,$$

and consider the optimization problem

$$\mathcal{Q}-\mathcal{LLS} \qquad \min_{x\in\mathbb{R}^n} \frac{1}{2}(Ax-b)^TQ(Ax-b).$$

(a) Give necessary and sufficient conditions under which the optimization problem $\mathcal{Q}-\mathcal{LLS}$ has a global optimal solution.
(b) If Q is positive definite, show that $\mathcal{Q}-\mathcal{LLS}$ is equivalent to a linear least squares problem.
(c) Give a necessary and sufficient condition under which $\mathcal{Q}-\mathcal{LLS}$ has a unique global optimal solution.

Solution Use the expansion

$$\frac{1}{2}(Ax-b)^TQ(Ax-b) = \frac{1}{2}x^TA^TQAx - (A^TQb)^Tx + \frac{1}{2}b^TQb$$

and apply known results.

(2) Determine whether the following matrices are positive definite, positive semi-definite, or neither by attempting to compute their Choleski factorizations.

$$(a)\ H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \qquad (b)\ H = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

$$(c)\ H = \begin{bmatrix} 5 & 2 & -1 \\ 2 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix} \qquad (d)\ H = \begin{bmatrix} 1 & 4 & 1 \\ 4 & 20 & 2 \\ 1 & 2 & 2 \end{bmatrix} .$$

Solution

(a) Positive definite by the determinant test. Gaussian elimination gives

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3/2 & 1 \\ 0 & 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 2/3 \\ 0 & 0 & 1 \end{bmatrix}$$

and so

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 0 & 2/3 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{3/2} & 0 \\ 0 & 0 & \sqrt{1/3} \end{bmatrix} .$$

(b) Obviously not positive definite since it has a -2 on the diagonal. Nonetheless, Gaussian elimination gives

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3/2 & 1 \\ 0 & 0 & -8/3 \end{bmatrix}$$

and so the last pivot being negative implies that H is not positive semi-definite.

(c) H is positive semi-definite. Gaussian elimination gives

$$\begin{bmatrix}1&0&0\\0&1&0\\0&3&1\end{bmatrix}\begin{bmatrix}1&0&0\\-2/5&1&0\\1/5&0&1\end{bmatrix}\begin{bmatrix}5&2&-1\\2&1&-1\\-1&-1&2\end{bmatrix}=\begin{bmatrix}5&2&-1\\0&1/5&-3/5\\0&0&0\end{bmatrix}=\begin{bmatrix}5&0&0\\0&1/5&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&2/5&-1/5\\0&1&-3\\0&0&0\end{bmatrix}$$

and so

$$L=\begin{bmatrix}1&0&0\\2/5&1&0\\-1/5&-3&0\end{bmatrix}\begin{bmatrix}\sqrt{5}&0&0\\0&\sqrt{1/5}&0\\0&0&0\end{bmatrix}.$$

(d) H is positive semi-definite. Gaussian elimination gives

$$\begin{bmatrix}1&0&0\\0&1&0\\0&1/2&1\end{bmatrix}\begin{bmatrix}1&0&0\\-4&1&0\\-1&0&1\end{bmatrix}\begin{bmatrix}1&4&1\\4&20&2\\1&2&2\end{bmatrix}=\begin{bmatrix}1&4&1\\0&4&-2\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&4&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&4&1\\0&1&-1/2\\0&0&0\end{bmatrix}$$

and so

$$L=\begin{bmatrix}1&0&0\\4&1&0\\1&-1/2&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&2&0\\0&0&0\end{bmatrix}.$$

(3) Consider the linearly constrained quadratic optimization problem

$$\mathcal{Q}(H,g,A,b)\qquad \begin{array}{ll}\text{minimize} & \frac{1}{2}x^THx+g^Tx\\ \text{subject to} & Ax=b\ ,\end{array}$$

where $H\in\mathbb{R}^{n\times n}$ is symmetric an positive definite and $A\in\mathbb{R}^{m\times n}$ has rank $(A)=m$.

(a) Write necessary and sufficient optimality conditions for this problem at a pair $(\bar{x},\bar{y})\in\mathbb{R}^n\times\mathbb{R}^m$ where $\bar{y}$ is an Lagrange multiplier vector.

Solution: $\begin{bmatrix}H&A^T\\A&0\end{bmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}-g\\b\end{pmatrix}$

(b) Solve the problem $\mathcal{Q}(H,g,A,b)$ with

$$H=\begin{bmatrix}1&1&0\\1&2&1\\0&1&3\end{bmatrix},\ g=(1,1,1)^T,\ b=(4,2)^T,\quad\text{and}\quad A=\begin{bmatrix}1&2&1\\1&0&1\end{bmatrix}.$$

Solution: Apply Gaussian elimination to the augmented matrix for the system given in (a),

$$\left[\begin{array}{ccccc|c}1&1&0&1&1&-1\\1&2&1&2&0&-1\\0&1&3&1&1&-1\\1&2&2&0&0&4\\1&0&1&0&0&2\end{array}\right],$$

to obtain the solution $\bar{x}=\frac{1}{2}(3,\ 2,\ 1)^T$ and $y=-\frac{1}{2}(5,\ 2)^T$.

(c) Solve the problem $\mathcal{Q}(H,g,A,b)$ with $g=0$, $m=1$, $b=\mu\in\mathbb{R}$, and $A=\mathbf{1}_n^T$, where $\mathbf{1}_n\in\mathbb{R}^n$ is the vector of all ones.

Solution: Again solve the system in part (a) to find that $x=\frac{\mu}{\mathbf{1}_n^TH\mathbf{1}_n}H^{-1}\mathbf{1}_n$ and $y=-\frac{\mu}{\mathbf{1}_n^TH\mathbf{1}_n}$.

(d) Show that the matrix $AH^{-1}A^T$ is invertible.

Solution: Since H is positive definite, it has a Cholesky factorization $H = LL^T$, where $L \in \mathbb{R}^{n\times n}$ is nonsingular. hence, $H^{-1} = L^{-T}L^{-1}$. Consequently, if $0 = AH^{-1}A^T x = AL^{-T}L^{-1}A^T x$, then $0 = x^T AL^{-T}L^{-1}A^T x = \|L^{-1}A^T x\|^2$, so $L^{-1}A^T x = 0$, or, equivalently, $A^T x = 0$ (just multipl through by L). But $\mathrm{Nul}\left(A^T\right) = 0$ since $\mathrm{rank}\,(A) = m$, so $x = 0$. Hence $AH^{-1}A^T$ is invertible.

(e) Show that $\bar{y} = -(AH^{-1}A^T)^{-1}(AH^{-1}g + b)$.

Solution: The system in part (a) tells us that $Hx+A^T y = -g$ and $Ax = b$. Multiplying the first expression through by H^{-1} gives $x = -H^{-1}(A^T y+g)$. Plugging this into $Ax = b$ gives $AH^{-1}A^T y = -(AH^{-1}g+b)$, or equivalently (by (d)) $y = -(AH^{-1}A^T)^{-1}(AH^{-1}g+b)$.

(f) Show that $\bar{x} = -[H^{-1} - H^{-1}A^T(AH^{-1}A^T)^{-1}AH^{-1}]g + H^{-1}A^T(AH^{-1}A^T)^{-1}b$.

Solution: In the proof of (e) we saw that $x = -H^{-1}(A^T y+g)$. Plugging the expression for y into this equation for x gives the result.

(g) Show that

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} [H^{-1} - H^{-1}A^T(AH^{-1}A^T)^{-1}AH^{-1}] & H^{-1}A^T(AH^{-1}A^T)^{-1} \\ (AH^{-1}A^T)^{-1}AH^{-1} & -(AH^{-1}A^T)^{-1} \end{bmatrix}.$$

Solution: Just do the matrix multiply.

(4) Prove Proposition 2.1 in Chapter 3.

Solution: Proposition 2.1 in Chapter 3 reads as follows:

Proposition 2.1 *Consider the two problems* $\mathcal{Q}(H, g, A, b)$ *and*

$$\widehat{\mathcal{Q}}(H, g, A, b) \qquad \textit{minimize } z \in \mathbb{R}^k \frac{1}{2} z^T \widehat{H} z + \hat{g}^T z \ ,$$

where $\widehat{H} := V^T H V$ *and* $\hat{g} := V^T(H\hat{x} + g)$ *with* $\hat{x}$ *any solution to* $Ax = b$ *and* V *being any matrix whose columns form a basis for the subspace* $\mathrm{Nul}\,(A)$. *Then the sets of optimal solutions to these problems are related as follows:*

$$\{\bar{x} \,|\bar{x} \text{ solves } \mathcal{Q}(H, g, A, b)\} = \left\{\hat{x} + V\bar{z} \,\middle|\, \bar{z} \text{ solves } \widehat{\mathcal{Q}}(H, g, A, b)\right\}.$$

Proof. Set $f(x) := \frac{1}{2}x^T Hx + g^T x$. Let $\bar{x}$ solve $\mathcal{Q}(H, g, A, b)$. Since $A\bar{x} = b$, there is a $\bar{z} \in \mathbb{R}^k$ such that $\bar{x} = \hat{x} + V\bar{z}$. In addition, since $A(\hat{x} + Vz) = b$ for any $x \in \mathbb{R}^k$, we have

$$f(\hat{x} + V\bar{z}) = f(\bar{x}) \leq f(\hat{x} + Vz).$$

Hence $\bar{z}$ solves $\widehat{\mathcal{Q}}(H, g, A, b)$.

Conversely, suppose $\bar{z}$ solves $\widehat{\mathcal{Q}}(H, g, A, b)$ and set $\bar{x} := \hat{x}+V\bar{z}$. For every $x \in \mathbb{R}^n$ satisfying $Ax = b$, there is a unique $z(x) \in \mathbb{R}^k$ such that $x = \hat{x} + Vz(x)$ since the columns of V form a basis for $\mathrm{Nul}\,(A)$. Hence, for every $x \in \mathbb{R}^n$,

$$f(\bar{x}) = f(\hat{x} + V\bar{z}) \leq f(\hat{x} + Vz(x)) = f(x),$$

and consequently $\bar{x}$ solves $\mathcal{Q}(H, g, A, b)$. □

(5) Prove Part (5) of Theorem 2.1 in Chapter 3 on the course notes.

Solution: Part (5) of Theorem 2.1 in Chapter 3 reads as follows:

Theorem 2.1 Part (5): *If either there exists $\bar{u} \in S$ such that $\bar{u}^T H \bar{u} < 0$ or there does not exist $\bar{x} \in \hat{x} + S$ such that $H\bar{x} + g \in S^\perp$ (or both), then*

$$-\infty = \inf_{x \in \hat{x}+S} \frac{1}{2} x^T H x + g^T x \ .$$

Proof. Set $f(x) = \frac{1}{2}x^T H x + g^T x$. If there is a $\bar{u} \in S$ for which $\bar{u}^T H \bar{u} < 0$, then

$$f(\hat{x} + t\bar{u}) = \frac{1}{2}(\hat{x} + t\bar{u})^T H(\hat{x} + t\bar{u}) + g^T(\hat{x} + t\bar{u}) = \frac{t^2}{2}\bar{u}^T H \bar{u} + t(H\hat{x} + g)^T \bar{u} + \frac{1}{2}\hat{x}^T H \hat{x},$$

where $(\hat{x} + t\bar{u}) \in \hat{x} + S$ for all $t \in \mathbb{R}$. Hence $f(\hat{x} + t\bar{u})^T H(\hat{x} + t\bar{u}) \downarrow -\infty$ as $t \uparrow +\infty$.

Next assume that H is positive semidefinite on S, but there is no $x \in \hat{x} + S$ such that $Hx + g \in S^\perp$. If $V \in \mathbb{R}^{n \times k}$ have columns that form a basis for S, where $k = \dim S$. Then our assumptions are equivalent to saying that $\widehat{H} := V^T H V$ is positive semidefinite and there is no solution to the equation $V^T(H(\hat{x} + Vz) + g) = 0$, i.e., $\hat{g} := V^T(H\hat{x} + g) \notin \operatorname{Ran}\left(\widehat{H}\right)$. Let $P \in \mathbb{R}^{k \times k}$ be the orthogonal projection onto $\operatorname{Ran}\left(\widehat{H}\right)$ so that $(I - P)$ is the orthogonal projection onto $\operatorname{Ran}\left(\widehat{H}\right)^\perp = \operatorname{Nul}\left(\widehat{H}\right)$. Set $\hat{g}_1 := P\hat{g}$ and $\hat{g}_2 := (I - P)\hat{g}$ so that $\hat{g} = \hat{g}_1 + \hat{g}_2$. Since $\hat{g} \notin \operatorname{Ran}\left(\widehat{H}\right)$, $\hat{g}_2 \neq 0$. Then, for all $t \in \mathbb{R}$, we have $x(t) := \hat{x} - tV(I - P)V^T(H\hat{x} + g) \in \hat{x} + S$ and

$$f(x(t)) = f(\hat{x}) + \frac{t^2}{2}\hat{g}_2^T \widehat{H} \hat{g}_2 - t\hat{g}^T \hat{g}_2 = -t\|\hat{g}_2\|^2.$$

Hence, $f(x(t)) \downarrow -\infty$ as $t \uparrow +\infty$, thereby proving the result. □

(6) Use Lemma 3.1 Part (1) to inductively show that the only matrix in $\mathcal{S}^n_+$ having zero diagonal is the zero matrix.

Solution: The proof follows by induction on the dimension n. The result is trivially true for $n = 1$, so suppose it is true for dimension $k = 1, \ldots, n-1$. We need to show it is true for dimension n. Let $H \in \mathcal{S}^n_+$ have zero diagonal and suppose

$$H = \begin{bmatrix} \widehat{H} & v \\ v^T & 0 \end{bmatrix}.$$

Clearly, $\widehat{H}$ is positive semidefinite with zero diagonal, and so, by the induction hypothesis, $\widehat{H} = 0$. Since $H \in \mathcal{S}^n_+$, Lemma 3.1 tells us that there is a vector $z \in \mathbb{R}^{n-1}$ such that $v = \widehat{H}z = 0z = 0$, which proves the result.

(7) Let $A \in \mathbb{R}^{m \times n}$ have rank $m < n$ and let $H \in \mathcal{S}^n$ be positive definite on $\operatorname{Nul}(A)$, i.e., $u^T H u > 0$, $\forall\, u \in \operatorname{Nul}(A)$. These hypotheses imply that AA^T is invertible. Define $A^\dagger := A^T(AA^T)^{-1}$. In this context the matrix $A^\dagger$ is called the Moore-Penrose pseudo inverse of A.

(a) Show that the dimension of the $\operatorname{Nul}(A)$ is $n - m$.

Solution: This is just the rank plus nullity theorem.

(b) Let $U \in \mathbb{R}^{n \times (n-m)}$ be *any* matrix whose columns form an orthonormal basis for $\operatorname{Nul}(A)$ and show that $I - UU^T = A^\dagger A$.

Solution: It was shown in clkass and the course notes that UU^T is the orthogonal projector onto $\operatorname{Nul}(A)$ and consequently $I - UU^T$ is the orthogonal projector onto $\operatorname{Nul}(A)^\perp = \operatorname{Ran}\left(A^T\right)$. Therefore, this question asks us to show that $A^\dagger A$ is the orthogonal projector onto $\text{Ran}\left(A^T\right)$. To this end, set $Q := A^\dagger A$ and note that $Q^2 = A^\dagger A A^\dagger A = A^T(AA^T)^{-1}AA^T(AA^T)^{-1}A = A^T(AA^T)^{-1}A = A^\dagger A = Q$ and $Q^T = (A^\dagger A)^T = A^T(AA^T)^{-1}A = Q$. Hence Q is the orthogonal projection onto its range. Thus, it remains only to show that $\text{Ran}\left(A^\dagger A\right) = \text{Ran}\left(A^T\right)$. Since $A^\dagger A = A^T(AA^T)^{-1}A$, it is clear that $\text{Ran}\left(A^\dagger A\right) \subset \text{Ran}\left(A^T\right)$. On the other hand, if $x \in \text{Ran}\left(A^T\right)$, then there is a u such that $x = A^T u$. Hence, $A^\dagger A x = A^\dagger A A^T u = A^T(AA^T)^{-1}AA^T u = A^T u = x$. Consequently, we have the reverse inclusion $\text{Ran}\left(A^T\right) \subset \text{Ran}\left(A^\dagger A\right)$ which establishes the result.

(c) Show that

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}^{-1} = \begin{pmatrix} U(U^T HU)^{-1}U^T & \left(I - U(U^T HU)^{-1}U^T H\right) A^\dagger \\ (A^\dagger)^T \left(I - HU(U^T HU)^{-1}U^T\right) & (A^\dagger)^T \left(HU(U^T HU)^{-1}U^T H - H\right) A^\dagger \end{pmatrix},$$

where $U \in \mathbb{R}^{n\times(n-p)}$ is any matrix whose columns form an orthonormal basis of $\ker A$.

Solution: Just use the two facts $(I - UU^T) = A^\dagger A$ and $AA^\dagger = I$ when multiplying the matrix given above by $\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}$ to show that you get the identity matrix.